A Classification of Matrices of Class Z

Miroslav Fiedler
Mathematics Institute
Czechoslovak Academy of Sciences
Žitná 25, Praha 1, 11567 Czechoslovakia

and

Thomas L. Markham*
Department of Mathematics
University of South Carolina
Columbia, South Carolina 29208

Submitted by Hans Schneider

ABSTRACT

We generalize the classes N_0 and F_0 studied by K. Fan, G. Johnson, and R. Smith. Schur complements and lattices are examined for matrices in these classes.

1. A GENERALIZATION OF N_0 AND F_0 MATRICES

Suppose A is a matrix over the field of real numbers. Throughout, we deal with $N \times N$ Z-matrices, which are matrices whose off-diagonal entries are nonpositive. We will use the notation of Fiedler and Pták [3] with regard to class K and class K_0, which are respectively the class of nonsingular M-matrices and the class of M-matrices.

Ky Fan [2] defined N-matrices to have the form

$$A = tI - B, \tag{1}$$

*The work of this author was done while visiting at the Mathematics Institute in Prague.

LINEAR ALGEBRA AND ITS APPLICATIONS 173:115–124 (1992) 115

655 Avenue of the Americas, New York, NY 10010 0024-3795/92/$5.00

where $B \geqslant 0$, and $\rho_{n-1}(B) < t < \rho(B)$, where $\rho(B)$ is the Perron eigenvalue of B and $\rho_{n-1}(B)$ is the maximum of the spectral radii of the $(n-1) \times (n-1)$ principal submatrices of B. He showed that N-matrices are Z-matrices with negative determinant and with proper principal submatrices belonging to K.

G. Johnson [5] extended Fan's definition to N_0-matrices. He required that N_0-matrices have the form given in (1) with $\rho_{n-1}(B) \leqslant t < \rho(B)$. He also studied matrices of the form (1) for $n \geqslant 3$, where $\rho_{n-2}(B) \leqslant t < \rho_{n-1}(B)$, where ρ_{n-2} is the maximum spectral radius of the $(n-2) \times (n-2)$ principal submatrices of B. Smith [6] called these matrices F_0-matrices, in honor of Fan. A further step was taken by Ying Chen [1], who also studied inverse F_0-matrices and inverse N_0-matrices.

We intend to generalize these ideas in the following manner.

Definition 1.1. Let L_s denote the class of real $n \times n$ matrices which have the form $A = tI - B$ where $B \geqslant 0$ and

$$\rho_s(B) \leqslant t < \rho_{s+1}(B), \tag{2}$$

where $\rho_s(B) = \max\{\rho(\hat{B}) : \hat{B} \text{ is a principal submatrix of } B \text{ of order } s\}$, for $s = 1, 2, \ldots, n$.

If $A \in L_s$, then we say the height of A in Z is s.

Using our notation, it is clear that $L_{n-1} = N_0$, $L_{n-2} = F_0$, $L_n = K_0$. For convenience, we let $\rho_s(B)\rho_{s+1}(B) = -\infty$ for $s = 0$, $\rho + 1(B) = \infty$ *for* $s = n$.

None of the classes $L_0, L_1, \ldots, L_n$ are void, as the following example demonstrates.

Example 1. Let J be the $n \times n$ matrix of all ones. Clearly, $\rho_k(J) = k$ for $k = 1, 2, \ldots, n$. If t satisfies $k \leqslant t < k+1$, then $tI - J \in L_k$ for $k = 1, \ldots, n$. Also, if $t < 1$, then $tI - J$ belongs to L_0.

If $B \geqslant 0$, it is well known that

$$\rho_1(B) \leqslant \rho_2(B) \leqslant \cdots \leqslant \rho_n(B) = \rho(B). \tag{3}$$

If B is strictly positive, the inequalities are strict. Further, if B is irreducible, then $\rho_{n-1}(B) < \rho_n(B) = \rho(B)$.

Too, we note that the class L_s is invariant under permutational similarity. The classes $L_0, L_1, \ldots, L_n$ form a decomposition of the class Z in the sense that if $A \in Z$, then A belongs to exactly one of these classes. This means, of course, that the height of A is well defined.

If A is a triangular matrix, then $A \in L_n$ if all diagonal entries of A are nonnegative. If at least one diagonal entry of A is negative, then A belongs to L_0. Thus, the triangular matrices belong only to the extreme classes.

The same is true also for diagonal matrices and for matrices obtained from triangular matrices by simultaneous permutations of rows and columns.

Next, we give an alternative definition of L_s, which is sometimes useful in our work.

DEFINITION 1.2. Suppose $1 \leqslant s \leqslant n - 1$. Let $\hat{L}_s$ denote the class of $n \times n$ Z-matrices which have the property that if $A \in Z$, all principal submatrices of A of order s belong to K_0, but there exists a principal submatrix of order $s + 1$ of A which is not in K_0. We define $\hat{L}_0 = L_0$ and $\hat{L}_n = L_n$.

THEOREM 1.3. *The classes L_s and $\hat{L}_s$ are identical for $s = 0, \ldots, n$.*

Proof. Let $s = 1, \ldots, n - 1$. ■

If $A \in L_s$, then $A = tI - B$, where $B \geqslant 0$ and $\rho_s(B) \leqslant t < \rho_{s+1}(B)$. Clearly, $A \in \hat{L}_s$.

If $A \in \hat{L}_s$, then since $A \in Z$, we write $A = tI - B$ with $B \geqslant 0$ and some scalar t. Since all principal submatrices of A of order s are in K_0, then we must have $t \geqslant \rho_s(B)$. If $t \geqslant \rho_{s+1}(B)$, then all submatrices of A of order $s + 1$ would be in K_0, which is a contradiction. Thus $\rho_s(B) \leqslant t < \rho_{s+1}(B)$, and A is in L_s.

COROLLARY 1.4. *If $A \in L_s$ and D is an $n \times n$ positive diagonal matrix, then DA and AD belong to L_s for $s = 0, 1, \ldots, n$.*

Proof. The proof follows immediately from Theorem 1.3, since a positive diagonal scaling of a matrix in K_0 remains in K_0, and if a principal submatrix of order $s + 1$ has a negative determinant, then a positive diagonal scaling of this submatrix also has a negative determinant. Thus both DA and AD belong to $\hat{L}_s$. ■

If A and B are $n \times n$ real matrices, we write as usual $A \leqslant B$ whenever $a_{ij} \leqslant b_{ij}$ for all pairs (i, j). The following is a generalization of G. Johnson's Theorem 2.10(i) in [5], and also shows the interplay of monotonicity with the class L_s.

THEOREM 1.5. *Let $A \in L_s$, $B \in L_t$ with $A \leqslant B$. Then*

(i) $s \leqslant t$, and

(ii) whenever a matrix C satisfies $A \leqslant C \leqslant B$, then $C \in L_q$, where $s \leqslant q \leqslant t$.

Proof. It suffices to prove the second part only. Since $C \in Z$, C belongs to some L_v, $0 \leqslant v \leqslant n$.

There exist nonnegative matrices P, Q, and R such that

$$\begin{aligned} A &= \lambda I - P, \qquad \rho_s(P) \leqslant \lambda < \rho_{s+1}(P), \\ B &= \lambda I - Q, \qquad \rho_t(Q) \leqslant \lambda < \rho_{t+1}(Q), \\ C &= \lambda I - R, \qquad \rho_v(R) \leqslant \lambda < \rho_{v+1}(R). \end{aligned} \tag{4}$$

Since $A \leqslant C \leqslant B$, the matrices P, Q, R satisfy

$$P \geqslant R \geqslant Q.$$

By a well-known property of nonnegative matrices,

$$\rho_u(P) \geqslant \rho_u(R) \geqslant \rho_u(Q), \qquad u = 0, \ldots, n.$$

But then it follows from (4) that

$$\rho_s(R) \leqslant \lambda < \rho_{t+1}(R).$$

By (3), $s \leqslant v \leqslant t$. ∎

LEMMA 1.6. *If $A = \lambda I - B$ and $\rho_s(B) = \rho_{s+1}(B) = \cdots = \rho_t(B)$ for some indices s, t, $s < t$, then A cannot belong to L_u for $s \leqslant u \leqslant t - 1$.*

Proof. Immediate. ∎

THEOREM 1.7. *Let $A \in L_s$ be reducible and $s < n$. Write $A = \lambda I - P$, where $p \geqslant 0$ has diagonal blocks (in the Frobenius form) $P_1, \ldots, P_k$. If $\rho(P) = \rho(P_j)$, then the order of P_j is greater than s.*

Proof. Let v be the order of P_j. Since $\rho_v(P) \geqslant \rho(P_j)$, we have, by (3),

$$\rho_v(P) = \rho_{v+1}(P) = \cdots = \rho_n(P).$$

Since $s < n$, it follows from Lemma 1.6 that $v > s$. ∎

COROLLARY 1.8 (G. Johnson [5]). *All matrices in L_{n-1} are irreducible.*

EXAMPLE 2. In the theorem, one cannot expect a stronger assertion, as the following example will illustrate. If

$$A = \begin{bmatrix} I_{n-k} & 0 \\ 0 & A_{22} \end{bmatrix},$$

where (a) A_{22} is $k \times k$, (b) $A_{22} = I - P$, where $P \geqslant 0$, and (c) A_{22} belongs to L_{k-1}, then A also belongs to L_{k-1}.

We shall also need the following example.

EXAMPLE 3. In Theorem 1.7, one cannot expect a stronger inequality for the order of P_j, as we now illustrate.

Let $J_{r,t}$ denote the $r \times t$ matrix of all ones. If $r = t$, we write J_r for simplicity. Now suppose that s and t are numbers such that $1 \leqslant s < t < n$.

Let

$$B = \begin{bmatrix} J_s & \varepsilon J_{s,t-s} & 0 \\ 0 & 0 & \varepsilon J_{t-s,n-t} \\ \varepsilon J_{n-t,s} & 0 & 0 \end{bmatrix},$$

where $\varepsilon > 0$ is chosen small. Let $A = \lambda I - B$. Then it is easily seen that $\rho_{s-1} = s - 1$ but $s = \rho_s(B) = \rho_{s+1}(B) = \cdots = \rho_t(B) < \rho_{t+1}(B)$. Therefore, the matrix $\lambda I - B$ belongs to L_s for $s - 1 \leqslant \lambda < s$, but $sI - B \in L_t$.

In the sequel, we shall investigate further properties of the classes L_s.

First, we denote by ∂L_s the set of matrices in Z for which equality is obtained in the left-hand-side inequality in (2):

$$\partial L_s = \{A \in Z : A = \lambda I - B, \lambda = \rho_s(B)\}, \qquad s = 1, \ldots, n.$$

We call ∂L_s the lower boundary of L_s.

It is clear that all the sets $L_s \diagdown \partial L_s$, $s = 1, \ldots, n-1$, are open sets in Z in the sense that whenever $A \in L_s \diagdown \partial L_s$, there exists a neighborhood N of A such that every matrix in $N \cap Z$ belongs to $L_s \diagdown \partial L_s$ as well. Also, L_0 is open. We can ask also about the closure $\overline{L}_s$ of L_s, i.e., the set consisting of all the limits of all convergent sequences of matrices belonging to L_s. It is clear that a matrix in L_j with $j < s$ cannot be in $\overline{L}_s$. Also, a matrix in $L_t \diagdown \partial L_t$ cannot belong to $\overline{L}_s$ for $t > s$. The following theorem shows that all other cases can occur.

THEOREM 1.9. *Let p, q be integers satisfying $0 \leqslant p < q \leqslant n$. Then*

$$\bar{L}_p \cap \partial L_q \neq \varnothing.$$

REMARK. Due to the previous observation, $\bar{L}_p \cap \partial L_q = \bar{L}_p \cap L_q$.

Proof. We shall distinguish five cases. In each case, we shall specify a matrix $V \in \partial L_q$ and a matrix U, and set

$$A_k = \frac{1}{k+1} U + V, \qquad k = 1, 2, \dots .$$

Thus $V = \lim_{k \to \infty} A_k$, and we always have that $A_k \in L_p$.

Case 1. $p = 0$, $q = n$. We set $V = 0$, $U = -I$. Clearly $A_k \in L_p$.
Case 2. $p = 0$, $q < n$. Define V as

$$V = \begin{array}{c} \left[\begin{array}{ccc} 0 & -e & 0 \\ 0 & 0 & -I \\ -e^T & 0 & 0 \end{array}\right] \begin{array}{l} n-q \\ q-1, \\ 1 \end{array} \\ \begin{array}{ccc} n-q & \;1\; & q-1 \end{array} \end{array}$$

where e is the vector of all ones, and $U = -I$.

Case 3. $p = 1$, $q \leqslant n$. Set V as in *case* 2, $U = I - J$, where J is a matrix of all ones.

Case 4. $p \geqslant 2$, $q = p + 1$. Set $V = (p+1)I - J$, $U = -I$.

Case 5. $p \geqslant 2$, $p + 1 < q \leqslant n$. Set $U = -I$ and $V = (p+1)I - B$, where B is the matrix from Example 3 with $s = p + 1$ and $t = q$. ∎

2. THE SCHUR COMPLEMENT

Suppose A is an $n \times n$ matrix partitioned as

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}. \tag{5}$$

If A_{11} is a square submatrix of A which is invertible, the Schur complement of A_{11} in A is

$$A/A_{11} := A_{22} - A_{21} A_{11}^{-1} A_{12}. \tag{6}$$

If A belongs to L_s, we will say that A is pure in L_s if every principal submatrix of order $s+1$ is negative.

We can now prove the following theorem.

THEOREM 2.1. *Suppose A is a matrix in L_s, and A_{11} is a $k \times k$ invertible submatrix of A, as in* (5), *where $k \leqslant s$.*

(i) *Then $A/A_{11} \in L_j$ for some $j \geqslant s-k$.*
(ii) *If A is pure in L_s, then A/A_{11} belongs to L_{s-k}.*

Proof. First, if A has order n, then A/A_{11} has order $n-k$. Let us index the rows and columns by $k+1,\dots,n$. If α is a sequence chosen from $k+1,\dots,n$, we let $A/A_{11}(\alpha)$ denote the principal submatrix of A/A_{11} whose rows and columns are indexed by α.

First, we note that $A/A_{11} \in Z$, since $A \in Z$ and $A_{11}^{-1} \geqslant 0$. It is well known that

$$\det A/A_{11}(\alpha_1,\dots,\alpha_k) = \frac{\det A(1,\dots,k,\alpha_1,\dots,\alpha_k)}{\det A_{11}}. \tag{7}$$

Hence all principal minors of A/A_{11} of order $\leqslant s-k$ are nonnegative. This proves the first assertion. If A is pure, there exists a principal minor of A/A_{11} of order $s-k+1$ which is negative. By Theorem 1.1, $A \in L_{s-k}$. ■

EXAMPLE 4. We return to the matrix of Example 3. For $A = \lambda I - B$ and $\lambda > s$, the Schur complement $A/\lambda I - J_s$ has the form

$$\begin{bmatrix} & 0_{t-s} & 0 \\ -\left(\dfrac{\epsilon}{\lambda} + \dfrac{\epsilon_s}{\lambda(\lambda - s)}\right) & (t-s)J_{n-t,t-s} & 0_{n-t} \end{bmatrix},$$

which clearly belongs to L_{n-s}. Although $A \in L_t$ where $t > s$, we have $A/A_{11} \in L_{n-s}$ for $n-s > n-t$. This shows that in the nonpure case, (i) cannot be improved.

3. LATTICES ASSIGNED TO A SQUARE MATRIX

In this section, we shall investigate the sign properties of the principal minors of a matrix in Z. For better understanding of these properties, we now introduce two definitions and some notation.

DEFINITION 3.1. Let A be a real $n \times n$ matrix. By the lattice of A, denoted by $\mathcal{L}(A)$, we mean the directed graph with 2^n points, each of which corresponds to exactly one subset M of the index set $N = \{1, 2, \dots, n\}$ and is assigned the sign $+1$, 0, or -1 of the principal minor $\det A(M)$ if $M \neq \emptyset$; the sign associated with $\emptyset$ is $+1$; the edges of $\mathcal{L}(A)$ are the ordered pairs (M_1, M_2) for which $M_2 \subset M_1$ and $|M_1| = |M_2| + 1$. We also call $|M|$ the level of the point M.

It is clear that the lattice of a matrix does not essentially change if we permute the rows and columns simultaneously.

We shall denote by T the point of the lattice corresponding to the set N, and B will be the point corresponding to the void set.

In the sequel, we denote by $\tilde{Z}_n$ the class of $n \times n$ matrices in Z whose principal submatrices of all orders are nonsingular.

DEFINITION 3.2. Let $A \in \tilde{Z}_n$. By the signed lattice of A, denoted by $\tilde{\mathcal{L}}(A)$, we mean the lattice $\mathcal{L}(A)$ defined earlier to which each edge (i, k) is assigned a sign $+1$ or -1 according as i and k have the same sign or opposite signs.

THEOREM 3.1. *Let $A \in Z$. Then $\mathcal{L}(A)$ has the following properties*:

(*i*) *If P is a point in $\mathcal{L}(A)$ such that one path from P to B consists of positively signed points only, then every path from P to B in $\mathcal{L}(A)$ has this property.*

(*ii*) *The height of A is equal either to n, or to the smallest level of a negative point (if such exists) diminished by one.*

Proof. The first property follows from the well-known property of matrices of class Z (cf. [3, Theorem 4.3]) that if there is one nested sequence of positive principal minors having length equal to the order of the matrix, then the matrix belongs to K and hence all nested sequences of principal minors of this length consist of positive elements.

The second assertion follows from the definitions. ∎

THEOREM 3.2. *Let A be a symmetric matrix in* $\tilde{Z}_n$.

(*i*) *All paths in* $\tilde{\mathscr{L}}(A)$ *from the point T to the point B contain the same number of positive edges (and, of course, of negative edges).*

(*ii*) *If* $A(M)$ *is a principal submatrix of A belonging to the class K, and* $M \subset N$ *with* $|M| = s$, *then the Schur complement* $A/A(M)$ *is a matrix in* $\tilde{Z}_{n-2}$, *and its signed lattice* $\tilde{\mathscr{L}}(A/A(M))$ *is obtained as the sublattice of* $\tilde{\mathscr{L}}(A)$ *induced (together with the signs) by* $\tilde{\mathscr{L}}(A)$ *on the set of those points which correspond to the subsets of indices of M.*

Proof. The first assertion follows by the Jacobi criterion: the inertia of A, In(A), is (π, ν, 0), where π is the number of coincidences of the signs plus one, and ν is the number of changes of the signs in one nested sequence of principal minors of A of length n (cf. [4, p. 272]).

The second assertion follows immediately from the formula (7). ■

THEOREM 3.3. *Let A be an* $n \times n$ *matrix in Z with height n (and thus in the class* K_0). *Then if* $\mathscr{L}(A)$ *contains a point M with sign zero, then every point from which there exists a path to M in* $\mathscr{L}(A)$ *has also sign zero.*

Proof. This follows from the fact (cf. [3, Theorems 5.5, 4.3]) that if A_{11} is a singular submatrix of $A \in K_0$, then every principal submatrix of A containing A_{11} is again singular. ■

The following notion could have importance for matrices in class Z.

DEFINITION 3.3. Let $A \in Z$. We say that A has positivity length l if l is the order of a maximal principal submatrix belonging to K_0.

REMARK. For $\mathscr{L}(A)$, l is the length of the longest nonnegative path in $\mathscr{L}(A)$ ending in B. Clearly, the matrices contained in proper classes L_s are characterized by the fact that their positivity length coincides with their height.

4. CONCLUDING REMARKS

CONJECTURE. If $\mathscr{L}$ is a lattice with 2^n points corresponding to all the subsets of $N = \{1, \dots, n\}$, and the signs of the points belong only to $\{0, 1\}$ and satisfy condition (i) of Theorem 3.1 and the condition of Theorem 3.3, then there exists an $n \times n$ matrix $A \in K_0$ such that $\mathscr{L} = \mathscr{L}(A)$.

The class of Z-matrices is closely related to the class of essentially nonnegative matrices of Varga [7]. Indeed, $A \in Z$ if and only if $-A$ is essentially nonnegative. Thus, all definitions and results of this paper have their counterpart for essentially nonnegative matrices.

REFERENCES

1 Ying Chen, Notes on F_0-matrices, *Linear Algebra Appl.* 142:167–172 (1991).
2 Ky Fan, Some matrix inequalities, *Abh. Math. Sem. Univ. Hamburg* 29:185–196 (1966).
3 M. Fiedler and V. Pták, On matrices with nonpositive off-diagonal elements and positive principal minors, *Czechoslovak Math. J.* 42:382–400 (1962).
4 F. R. Gantmacher, *Theory of Matrices*, Nauka, Moscow, 1953.
5 G. Johnson, A generalization of N-matrices, *Linear Algebra Appl.* 48:201–217 (1982).
6 R. L. Smith, Bounds on the spectrum of nonnegative matrices and certain Z-matrices, *Linear Algebra Appl.*, 129:13–28 (1990).
7 R. S. Varga, *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, N.J., 1962.

Received 27 December 1990; final manuscript accepted 9 July 1991